浙江大學
ZHEJIANG UNIVERSITY

Restricted Pooling

Introduction

- To promoting the problem of benign or malignant diagnosis of pulmonary nodule, a novel pooling strategy is presented.

- This method is named as restricted pooling.

- It focuses on regularization of networks, which enhances the capability of robustness and generalization.

Introduction

- Generally, neural networks with high capacity, which means owning massive learnable parameters or lots of layers in topology, can minimize the cost function quickly and largely in the training session.

- However, there is a remarkable contrast in the test session since the cost function's value reduces much more slower.

Introduction

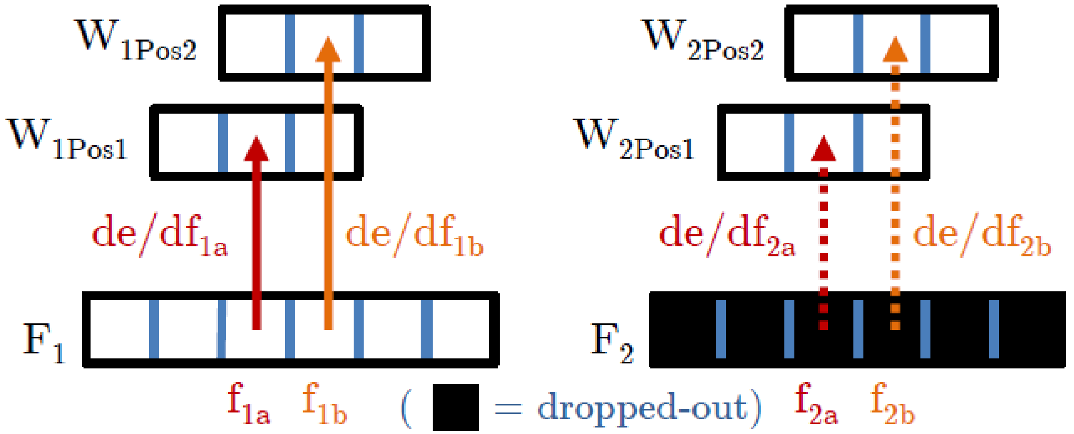
- The main insight of our work is that neural networks learn massive low-relevancy features in the training process and put relatively heavy weight on them when they are inferencing.

- These features make sense in specific circumstance, i.e., the dataset used for training. In other words, they are not necessary for a correct inference.

- More supervision should be given in the training session of networks, features used for training require to be selected cautiously.

Related work

- Efficient Object Localization Using Convolutional Networks. (Spatial Dropout)



- The dropout value is across the entire feature map.

Related work

- DropBlock: A regularization method for convolutional networks

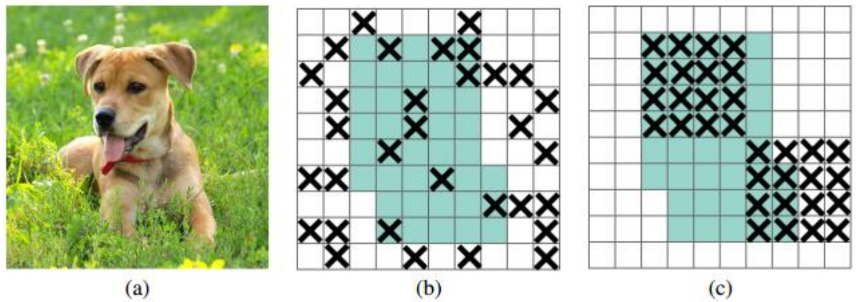
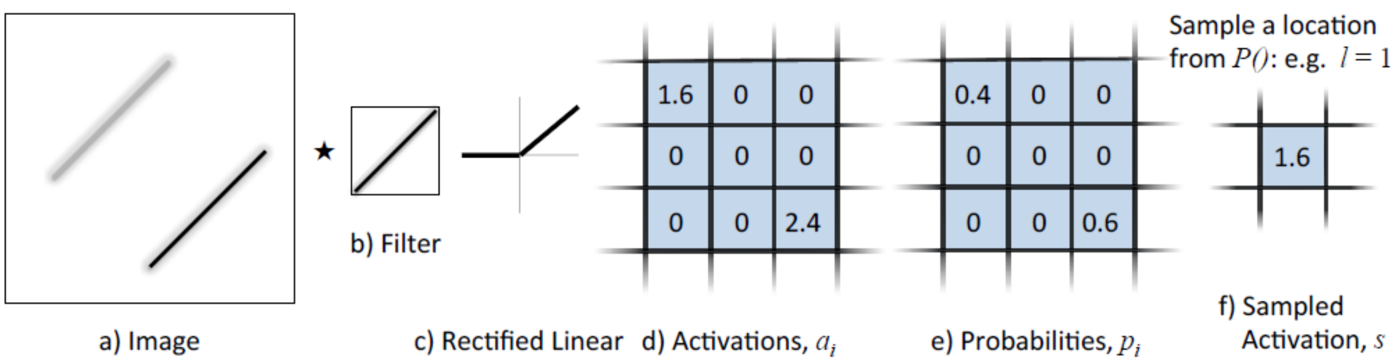


Figure 1: (a) input image to a convolutional neural network. The green regions in (b) and (c) include the activation units which contain semantic information in the input image. Dropping out activations at random is not effective in removing semantic information because nearby activations contain closely related information. Instead, dropping continuous regions can remove certain semantic information (e.g., head or feet) and consequently enforcing remaining units to learn features for classifying input image.

- Dropout commonly is not effective for convolutional neural networks, because the features of feature-map are high-relevancy in space.
- ImageNet: ResNet-50 76.51|78.13
- COCO: RetinaNet 36.8|38.4.

Related Work

- Stochastic Pooling for Regularization of Deep Convolutional Neural Networks



- Assigning probability according the value of feature-maps activated by ReLU.
- CIFAR-10: 19.24|15.13 MINIST: 0.55|0.47
- CIFAR-100: 47.77| 42.51

Method

- Novel pooling method is proposed basing on the insight we mentioned in the introduction.
- Restricted pooling is a structured method contains a estimator to evaluate the relative importance of feature-maps' different parts.
- Under the supervision of restricted pooling, those low-relevancy part will be "dropout" in the pooling process.

Method

- In the two-dimensional case, restricted pooling firstly divide feature-maps into N patched with same size: $P_{11}, P_{12}, ..., P_{1\sqrt{N}}, ..., P_{\sqrt{N}1}, P_{\sqrt{N}2}, ..., P_{\sqrt{N}\sqrt{N}}$.

- The estimator $E(\cdot)$ take all P_{ij} as inputs, for each it finally calculates a feature score S_{ij} to measure P_{ij} the relevancy of features contained in the patch to the task.

- N/d^2 patches finally selected according to the feature score, where d denotes the down scale factor.

Method

$$O_k = \Sigma_{i,j} select(S_{ij} - max^k_{i,j}\{S_{ij}\}) \odot P_{ij}$$

- where O_k denotes the kth selected patch, $max^k_{i,j}\{S_{ij}\}$ means the kth largest score of set $\{S_{ij}\}$, $\odot$ is a element-wise multiplication. select() indicates a discontinuous function defined as below:

$$select(x) = \begin{cases} 1 & |x| < eps \\ 0 & |x| \geq eps \end{cases}$$

- eps equals to 1e-5 to enhance numerical stability, after additional definition, select() turns to be differentiable:

Method

$$\frac{\mathrm{d}select(x)}{\mathrm{d}x} = \begin{cases} 0 & |x| < eps \\ 1 & x \leq -eps \\ -1 & x \geq eps \end{cases}$$

- The $\bar{N}/d^2$ outputs of restricted pooling are concatenated as a new feature-map in the end, which is the final pooling result.

Experiments

- Restricted pooling is firstly implemented on LIDC-IDRI dataset， this dataset contains 1361 benign lung nodule samples， 640 malignant ones , and 636 cases are not sure for theirs label.

- Restricted pooling shows powerful regularization, in the task of classification, the network performs equally accurate on both test and training datasets. To preventing underfitting, it is only applied on shallow layers of networks in some cases.

Experiments

- VGG16, LIDC-IDRI

Network	Accuracy	AUC
VGG16	86.61	89.41
VGG16+rpd4	88.12	91.44
VGG16+rpd6	87.54	91.29

- ResNet50,LIDC-IDRI

Network	Accuracy	AUC
ResNet50	86.46	
ResNet50+rpd4(all)	84.96	
VGG16+rpd4(shallow)	88.44	

Thanks!